ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Applied Logic 5 (2007) 1–2

JOURNAL OF APPLIED LOGIC

www.elsevier.com/locate/jal

Editorial

Questions and Answers: Theoretical and Applied Perspectives

This special issue is the outcome of dissemination activities carried out within the Network of Excellence in Computational Logic (CoLogNET http://www.cologne.org). In particular, it results from an exploration of the common grounds between the "Logic and Natural Language Processing" Area of CoLogNET and ElsNET (Network of Excellence in Human Language Technologies, http://www.elsnet.org/) that has been the aim of the annual CoLogNET-ElsNET symposia organized from 2001 to 2004.

Motivated by the broad interest that question answering (QA) is receiving in such fields as computational linguistics, computational logic, databases, formal semantics, and information retrieval, and inspired by the success of the second CoLogNET-ElsNET Symposium on "Questions and Answers: Theoretical and Applied Perspectives", we were led to prepare a special issue of the *Journal of Applied Logic* that addresses both the logical foundations underlying QA and technological implications for QA systems.

Abstracting from the different approaches and angles, it is clear that two central concerns underlying current research in QA are (1) integrating current open-domain QA (based on free "unstructured" text) with its more mature sister, natural language front ends to database systems (NLDBS), and (2) issues of what should be taken as "answer". Both aspects have been carefully discussed by Spärck-Jones [1]. In this paper, the author invites more hospitable and flexible research on QA that learns from its own history and its interaction with other disciplines and that attributes more importance to the problem of ranking multiple candidate answers. The papers collected in this special issue can be seen as responses to this invitation.

The interaction between open-domain QA and NLDBS is the focus of both the papers by Badia and by Frank et al. included in this special issue. In the former, Badia proposes a query language with generalized quantifiers to help bridge the gap between the two fields. The developed query language is meant to access information using a single formal language independently of the source from which the answer is going to be retrieved. The need for such a language comes from the observation that QA and NLDBS are converging. For instance, in the database community semistructured and unstructured data are acquiring more and more importance. Hence, the need for query languages that are more flexible than SQL and the appeal to tools and techniques from information retrieval.

A similar desire to unify QA access to information sources with different degrees of structuring is behind the hybrid architecture described in the paper by Frank et al. They present their approach for domain-restricted question answering from structured knowledge sources and multilingual data (QUETAL). The distinctive characteristics of this system are its hybrid architecture integrating high-quality NLP tools, the focus on linguistic analysis of questions, and its use of ontologies to interface between question analysis, answer extraction and knowledge engineering. It also contains several abstraction layers that reduce the complexity of the mapping rules to database concepts and guarantee portability across languages and across different target knowledge bases. (The paper by Moldovan et al. provides another example of hybrid architecture for open-domain QA.)

The second concern addressed by the papers in this special issue is what constitutes an answer and how answer relevance can be measured and ranked. Different proposals have been put forward.

Research on open domain QA systems has generally avoided deep logical approaches, preferring instead a combination of information retrieval and extraction and statistical approaches. Two of the papers in this special issue describe successful open domain QA system that take advantage of logic based approaches. Moldovan et al. use automated reasoning to enhance their QA system COGEX and use a prover to check whether a given question is entailed

1570-8683/$ – see front matter
doi:10.1016/j.jal.2005.12.008

by a retrieved candidate answer together with a large set of different kinds of axioms (WordNet, linguistic, ontology and temporal axioms). Their proof system also provides information for assigning confidence scores to the candidate answers and a basis for generating answer justification.

A logic based theorem prover for QA is also the focus of the paper by Burhnas and Shapiro. They propose to broaden the notion of answer to every reasoning step of a resolution theorem prover so as to avoid overlooking some information that could be relevant for a questioner, and to classify the answers into: specific, generic and hypothetical answers. These classes are shown to be relevant for QA systems and crucial to avoiding misleading answers (i.e., correct answers that nevertheless lead the questioner to draw incorrect conclusions based on some incorrect expectation).

An alternative solution to these deep logical analyses based on classical logic is given by De Boni who refers to the notion of "relevance" to judge correct answers and rank them. This approach allows to extract (relevant) answers even when the information sources are incomplete, inconsistent or errorful. Logical relevance is computed through the use of relaxation rules, with relevance defined in terms of the effort required to prove an answer by relaxing constraints. (Relaxing constraints are also used in the back-off module of Moldovan et al.'s COGEX system.)

The great diversity of candidate answers that can be relevant to a questioner is exploited in an original way in the paper by Dalmas and Webber. Although not the only paper in this special issue concerned with retrieved answers and their relevance and re-ranking, Dalmas and Webber follow a Model-View-Controller (MVC) design strategy, treating answers as structured objects contained in a model and retrieved by a strategy to build a view. The paper presents interesting results both for answer re-ranking and answer rendering, the latter obtained by applying an MVC design strategy to multi-document summarization and information fusion. In this work, candidate answers are considered as allies rather than competitors and even incorrect answers are shown to help reaching the required information due to their relation to the correct answer topic.

The problem of re-ranking is also the focus of the paper by Schlobach et al. As opposed to e.g., Moldovan et al., who use a very knowledge-intensive selection process and extensive lexical resources, these authors use data-driven scoring methods. They present a detailed comparison and experimental evaluation of different type checkers given the task of filtering out incorrect answers. They show how knowledge-poor type checking methods can achieve the same overall performance as knowledge-intensive type checking, through the use of simple web mining efforts and information redundancy, and they report on experiments that evaluate the impact of type checking methods on the performance of their QA system (QUARTZ). Finally, they provide an extensive error analysis explaining the meaning of their results and highlighting issues for further research.

As a last word, we would like to thank the other members of the Program Committee of the second CoLogNET-ElsNET Symposium on "Questions and Answers: Theoretical and Applied Perspectives"—Enrico Franconi, Michael Moortgat, Maarten de Rijke and Hans Uszkoreit—for their valuable help in selecting the papers submitted to this special issue and all the anonymous referees that have helped in putting it together. We hope that this collection will be of interest to the community and a stimulus for further collaboration between linguists, logicians, computer scientists, and cognitive scientists to address the challenges raised by "question and answers". Finally, we would like to thank CoLogNET and ELsNET for the financial support for the symposium that gave rise to this special issue.

Reference

[1] K. Spärck-Jones, Is question answering a rational task? in: R. Bernardi, M. Moortgat (Eds.), 2nd CoLogNET-ElsNET Symposium. Questions and Answers: Theoretical and Applied Perspectives, 2003.

Raffaella Bernardi
E-mail address: bernardi@inf.unibz.it

Bonnie Webber
E-mail address: bonnie@inf.ed.ac.uk

Available online 26 January 2006